Cybernetics, Computationalism, and Enactivism: Lessons from a Forgotten Past

By: Joe Dewhurst

In this paper I will identify the shared ancestry of two opposing paradigms in contemporary cognitive science, and argue that we can draw on this ancestry to help reconcile the two paradigms. The two paradigms are the computational and enactive approaches to cognitive science, and their shared ancestry is the cybernetics research program of the 1940s and 50s. I will first present the two paradigms and describe their contemporary opposition to one another, before exploring the cybernetic foundation that they both share. This, I will argue, contains much of the intellectual resources required in order to enable a fruitful reconciliation between the two paradigms. I will not attempt to fully articulate that reconciliation here, but I will conclude by suggesting a few promising avenues for future research. The computationalist paradigm that is currently dominant in the cognitive sciences, and especially in cognitive neuroscience, takes as its foundational assumption the idea that the primary function of the brain and nervous system is to perform computations (Piccinini 2012). Explanations of cognitive phenomena are given in terms of the performance of computations, and whilst there is some disagreement about the form that these computations take, it is agreed that the positing of computational mechanisms is a central feature of cognitive scientific explanation. Paradigmatic examples of this approach include Marr's theory of vision (1982), Fodor's language of thought hypothesis (1975), and more recently, connectionist and Bayesian interpretations of cognitive computation.

Since Varela, Thompson, & Rosch (1991), the minority opposition to the computationalist paradigm has typically been referred to as "enactivism". The term really covers a family of distinct but related views (Villabos & Ward 2015), and here I extend its usage to include dynamicist and embodied approaches more generally. What these approaches all have in common is that they reject the primacy of computational explanation in favour of explanations that emphasise the role played by non-computational factors such as the environment (both physical and social) and the body. Some of these approaches also deny that the brain or nervous system computes at all, although, as I will try to show, this is not essential to the definition of enactivism (just as a denial of bodily or environmental contributions to cognition is not essential to the definition of computationalism).

Both the computational and enactivist approaches to cognitive science owe their existence to the cybernetics research program of the 1940s and 50s. This grew out of interdisciplinary connections made during the war, and solidified in and around the Macy Conferences from 1946-53 (Dupuy 2009: 71- 5). At the heart of the program was the computational modelling of the brain pioneered by McCulloch & Pitts (1943) and popularised by Von Neumann (1958), but it also included insights such as the application of control systems analysis to living organisms (Wiener 1948), circularity and feedback as a source of natural teleology (Rosenblueth, Wiener, & Bigelow 1943), and the idea of biological homeostasis embodied in Walter's various robotic creations (1950, 1951). This era was the dawn of cognitive science as we know it today, and one of its most interesting features is the fruitful collaboration between what would go on to become the core principles of computationalism and enactivism: internal computational structure on the one hand and embodied interaction with the environment on the other.

This collaboration was possible because computation was not yet conceived of as a primarily representational phenomenon. Following Turing (1936), there was a lively interest in the abstract philosophical and mathematical understanding of computation, but for many working in cybernetics computation' was still primarily an engineering notion, and a computer was a very real and very physical entity. This allowed for experiments such as those conducted by Walter, which were more concerned with the interaction between a simple computational system and its environment than with the cognitive capacities of computation in the abstract (Walter 1953; Pickering 2010: 37-90). It was not until the philosophical work of Putnam and Fodor in the 60s, 70s, and 80s that computationalism came to be primarily concerned with an abstract and representational computational theory of mind (see e.g. 1967a, 1967b; Fodor 1975, 1981). Perhaps not coincidentally, it was also at this point that cybernetics as a distinct research program began to collapse, due in no small part to the pressures of carrying out interdisciplinary research in an increasingly compartmentalised academic environment (Pickering 2010).

There need not be any fundamental dichotomy between the enactivist and computationalist programs in cognitive science. Recent work in philosophy of computation has led to the emergence of distinct theories of physical computation, and in particular Piccinini's mechanistic account provides a promising non-representational foundation for reconciliation with the enactivist program (Piccinini 2007, 2015). There are some remaining issues that will need to be resolved before such this reconciliation can take place, such as whether mechanistic explanation can be understood without any notion of teleology (Dewhurst forthcoming; Piccinini 2015: 100-17), and how the enactive notion of functional closure can best be applied to computational systems (Dewhurst 2016) but there is no in principle reason why this should not be possible. More positively, a reconciled approach to cognitive science would have a lot to offer to precisely those issues that cybernetics originally focused on such as the interaction between perception and action and the emergence of natural teleology from environmental and social feedback. These are issues that remain central to contemporary cognitive science, as demonstrated by the recent emergence of the action oriented predictive coding framework (Clark 2016) and developments in social cognition that emphasise the role of participatory sense-making (De Jaegher & Di Paolo 2007). Here there are important lessons to be learnt from our forgotten cybernetic past, which not only provides a demonstration that computation and environmental interaction need not be in opposition, but also contains intellectual and explanatory tools that can be reclaimed and put to good use. Without returning to this history and considering how we got to where are today, we may never truly be able to move forwards.